

# Validating the Automated Assessment of Participation and of Collaboration in Chat Conversations

Mihai Dascalu<sup>1</sup>, Stefan Trausan-Matu<sup>1</sup>, Philippe Dessus<sup>2</sup>

<sup>1</sup> Politehnica University of Bucharest, Computer Science Department, Romania  
mihai.dascalu@cs.pub.ro, stefan.trausan@cs.pub.ro

<sup>2</sup> LSE, Univ. Grenoble Alpes, France  
philippe.dessus@upmf-grenoble.fr

**Abstract.** As Computer Supported Collaborative Learning (CSCL) gains a broader usage as a viable alternative to classic educational scenarios, the need for automated tools capable of supporting tutors in the time consuming process of analyzing conversations becomes more stringent. Moreover, in order to fully explore the benefits of such scenarios, a clear demarcation must be made between *participation* or active involvement, and *collaboration* that presumes the intertwining of ideas or points of view with other participants. Therefore, starting from a cohesion-based model of the discourse, we propose two computational models for assessing collaboration and participation. The first model is based on the cohesion graph and can be perceived as a longitudinal analysis of the ongoing conversation, thus accounting for participation from a social knowledge-building perspective. In the second approach, collaboration is regarded from a dialogical perspective as the intertwining or overlap of voices pertaining to different speakers, therefore enabling a transversal analysis of subsequent discussion slices.

**Keywords:** Computer Supported Collaborative Learning, Cohesion-based Discourse Analysis, Dialogism, Participation Assessment, Collaboration Evaluation.

## 1 Introduction

Computer Supported Collaborative Learning (CSCL) gains a broader usage in several newest educational settings, like MOOCs or collaborative serious games, as a viable alternative to classic educational scenarios. The need for automated tools capable of supporting all their actors in the time consuming process of analyzing conversations becomes more stringent. Chat conversations or forums became the place where knowledge is collaboratively built and shared [1] and there is a complex intertwining between collective and individual learning processes that is worth analyzing [2].

Shortly put, two complementary analysis approaches compete. The first one is *structural*, uses Social Network Analyses and stems from group dynamics to unveil relationships between individuals to sketch networks of collaboration [3]. The second approach is *dialogical*, has roots in discourse theories [4] and uses Natural Language

Processing techniques to analyze the semantic cohesion of textual utterances (e.g., sentences or paragraphs).

After devising several systems inspired from the dialogical approach [5] and using a cohesion-based model of the discourse as underlying structure [6], we propose computational models for assessing participation and collaboration. Within our approach, *participation* is regarded as cumulative qualitative utterance scores and is modeled through the interaction graph presented in the second section. Section three introduces two computational models for assessing *collaboration*. The first one is based on the cohesion graph [7] and can be perceived as a longitudinal analysis of the ongoing conversation, thus accounting for participation from a social knowledge-building perspective. In the second model, collaboration is regarded from dialogism as the intertwining or overlap of voices pertaining to different speakers, therefore enabling a transversal analysis of subsequent discussion slices. This paper is the occasion to present in the fourth section the results of a large-scale validation by comparing the outputs of our system with human evaluations.

## 2 Participation Assessment

Measuring participation in virtual groups and communities on the web communicating through chats, forums or different types of social networking was performed in the *structural approach* by considering the number of emitted posts or utterances and by using several social networks metrics like centrality (number of links to other nodes), betweenness (nodes that, if eliminated would highly reduce or eliminate communication among other participants) [8] or page-rank derived formulas [9] in the interaction graph with users as nodes and posts as arcs [9]. Sometimes arcs have weights computed in different ways, from the simplest number of posts to more complicated metrics, considering the language content of the messages, like in our *dialogical approach*, which will be presented below.

The assessment of participation of each student in CSCL chats has some differences from the cases of forums or other social networking due to the small number of participants (typical examples are 3 to 7 students) and the large number of exchanged utterances. In this case, due to the fact that for chat conversations we are dealing in most cases with a complete graph, betweenness score for all nodes is 0. Centrality also is not a very significant discriminant: only participants with very low number of emitted utterances are not central.

In our approach, we are taking a perspective based on natural language processing of the *content of utterances*, considering the *topics* that were supposed to be discussed (for example, stated by the teacher in a CSCL homework) and focusing on discourse analysis. The latter's defining feature is *cohesion* and our approach is fundamentally based on it. From a computational point of view, cohesion is computed as a combination of semantic distances in ontologies, semantic similarity from Latent Semantic Analysis vector spaces, and Latent Dirichlet Analysis topic models [7]. Starting from this aggregated similarity function, a multi-layered cohesion graph is built [10] that models through cohesive links the dependencies between the key

elements of the analysis: the whole conversation, participants' utterances and sentences for longer posts. The previous links can be either *explicit* if participants marked the dependencies within the user interface, *enforced* for hierarchical links and adjacent analysis elements or *implicit*, if cohesion exceeds a threshold value [10].

In terms of participation, we start with the identification of discussion topics for each participant for pinpointing out if the needed concepts were covered. One of the most important metrics is the *utterance score* that, shortly put, represents the overall topics coverage augmented through cohesion with inter-linked analysis elements [10]. In this aim, an *interaction graph* is built with participants as nodes and the weight of links equal to the sum of utterances scores multiplied by the cohesion with the inter-linked analysis elements [10].

### 3 Collaboration Evaluation

In order to thoroughly assess collaboration, we have proposed two computational models. The first model [6] represents a refinement of the gain-based collaboration assessment [11] and takes full advantage of the cohesion graph [12]. The second is a novel approach that evaluates collaboration as an intertwining or overlap of voices pertaining to different speakers. The main difference between the two is that the first focuses on the ongoing conversations, therefore on its longitudinal dimension, whereas the later considers subsequent slices of the conversation, the synergy of voices, in other words the transversal dimension.

#### 3.1 Social Knowledge-Building Model

The actual information transfer through cohesive links from the cohesion graph can be split between a personal and a social knowledge-building process [1, 13, 14] at utterance level. Firstly, a *personal dimension* emerges by considering utterances with the same speaker, therefore modeling a kind of inner voice or continuation of the discourse. Secondly, inter-changed utterances with different speakers define a *social perspective* that models collaboration as a cumulative effect. Our model is similar to some extent to the gain-based collaboration model [11] and marks a transition towards Stahl's model of collaborative knowledge-building [1] by representing a conversation thread as a multi-layered cohesion graph.

The continuation of ideas or explicitly referencing utterances of the same speaker builds an inner dialogue or personal knowledge, whereas the social perspective measures the interaction with other participants, encourages idea sharing and fosters creativity for working in groups [15], thus enabling a truly collaborative discussion. Moreover, personal knowledge building addresses individual voices (participant voices or implicit/alien voices covering the same speaker), while social knowledge building, derived from explicit dialog (that by definition is between at least two entities), sustains collaboration and highlights external voices.

### 3.2 Dialogical Voice Inter-Animation Model

In order to achieve genuine collaboration, the conversation must contain threads of utterances integrating key concepts ('voices', in the musical polyphonic sense [16]) that inter-animate in a similar way to counterpoints in polyphonic fugues. Voices are present in utterances from multiple participants of the conversation. In order to obtain an operationalization, a shift of perspective is required from voices, computed as semantic chains of related concepts, towards an individual participant. As collaboration is centered on multiple participants, a split of each voice into multiple viewpoints pertaining to different participants is required. A viewpoint consists of a link between the concepts pertaining to a voice and a participant, through their explicit use within one's interventions in the ongoing conversation. We opted to present this split in terms of implicit (alien) voices [17]. Moreover, this split presentation of semantic chains per participant is useful for observing each speaker's coverage and distribution of dominant concepts throughout the discussion.

In addition, in order to identify the voice overlaps now pertaining to different participants, we changed from an ongoing longitudinal analysis of the discourse, presented in the previous subsection, to a transversal analysis of a context consisting of several adjacent utterances. We use a cumulated value of Pointwise Mutual Information (PMI) obtained from all possible pairs of voices pertaining to different participants (different viewpoints), within subsequent contexts of the analysis (within our implementation we used a sliding window of 5 interventions in order to model the local context of each voice occurrence). From an individual point of view, each participant's overall collaboration can be seen as the cumulated mutual information between his personal viewpoints and all other participant viewpoints. Therefore, by comparing individual voice distributions that span throughout the conversation, collaboration emerges from the overlap of voices pertaining to different participants.

## 4 Participation and Collaboration Validation

The validation experiments focused on the assessment of 10 chat conversations that took place in an academic environment in which Computer Science students from the 4<sup>th</sup> year undergoing the Human-Computer Interaction course at our university debated on the advantages and disadvantages of CSCL technologies. Each conversation involved 4 or 5 participants who each had to support a given technology (e.g., chat, blog, wiki, forum or Google Wave) in specific use case scenarios during the first phase of the discussion, later on proposing an integrated alternative that would encompass the previously presented advantages. The 10 conversations were manually selected from a 10 times larger corpus of chats.

Afterwards, 76 4<sup>th</sup> year undergraduate students following the same course, but from a different generation, and 34 1<sup>st</sup> year master students attending the Adaptive and Collaborative Systems course were each asked to manually annotate 3 chat conversations. We opted to distribute the evaluation of each conversation due to the high amount of time it takes to manually assess a single discussion (on average, users reported 1.5 to 4 hours for a deep understanding) [18]. In the end, we had on average

33 annotations per conversation and the overall results indicated a reliable automatic evaluation of both participation and collaboration. We validated the machine vs. human agreement by firstly computing intra-class correlations between raters for each chat ( $\text{avg ICC}_{\text{participation}} = .97$ ;  $\text{avg ICC}_{\text{collaboration}} = .90$ ) and, secondly, as these correlations were all very high indicating very few disagreements between raters, non-parametric correlations ( $\text{avg } Rho_{\text{participation}} = .84$ ;  $\text{avg } Rho_{\text{collaboration}} = .74$ ) were determined between machine vs. human mean ratings for each chat.

## 5 Conclusions and Future Research Directions

Starting from a dialogic model of discourse centered on cohesion, we thoroughly validated our system in terms of analyzing chat participants' involvement and collaboration, the later employing a longitudinal model based on social knowledge-building and a different transversal model based on voice inter-animation. Moreover, as the validations proved the accuracy of the models built on dialogism, we can state that the proposed methods emphasize the dialogical perspective of collaboration in CSCL environments.

In addition, the analyses performed in this paper have a very broad spectrum of applications, extending from utterance cohesion towards group cohesion rooted in collaboration. Beyond the rather simple visualization of individual and collective involvement, our developed system is also well-suited to enable students to self-regulate their learning.

## Acknowledgements

We would like to thank the students of University “Politehnica” of Bucharest who participated in our experiments. This research was partially supported by the 264207 ERIC–Empowering Romanian Research on Intelligent Information Technologies/FP7-REGPOT-2010-1 project.

## References

1. Stahl, G.: Group cognition. Computer support for building collaborative knowledge. MIT Press, Cambridge, MA (2006)
2. Cress, U.: Mass collaboration and learning. In: Luckin, R., Puntambekar, S., Goodyear, P., Grabowski, B., Underwood, J., Winters, N. (eds.) Handbook of design in educational technology, pp. 416–424. Routledge, New York (2013)
3. Reffay, C., Martinez-Mones, A.: Basic concepts and techniques in social network analysis. In: Luckin, R., Puntambekar, S., Goodyear, P., Grabowski, B., Underwood, J., Winters, N.

- (eds.) *Handbook of design in educational technology*, pp. 448–456. Routledge, New York (2013)
4. Bakhtin, M.M.: *The dialogic imagination: Four essays*. The University of Texas Press, Austin and London (1981)
  5. Dascalu, M., Rebedea, T., Trausan-Matu, S., Armit, G.: PolyCAFe: Collaboration and Utterance Assessment for Online CSCL Conversations. In: CSCL 2011, Vol. 2, pp. 781–785. ISLS, Hong Kong, China (2011)
  6. Dascalu, M., Trausan-Matu, S., Dessus, P.: Cohesion-based Analysis of CSCL Conversations: Holistic and Individual Perspectives. In: CSCL 2013, Vol. 1, pp. 145–152. ISLS, Madison, USA (2013)
  7. Dascalu, M., Dessus, P., Trausan-Matu, S., Bianco, M., Nardy, A.: ReaderBench, an Environment for Analyzing Text Complexity and Reading Strategies. In: AIED 2013, Vol. LNCS 7926, pp. 379–388. Springer, Memphis, USA (2013)
  8. Brandes, U.: A Faster Algorithm for Betweenness Centrality. *Journal of Mathematical Sociology*, 25(2), 163–177 (2001)
  9. Dascalu, M., Chioasca, E.V., Trausan-Matu, S.: ASAP – An Advanced System for Assessing Chat Participants. In: AIMS 2008, Vol. LNCS 5253, pp. 58–68. Springer, Varna, Bulgaria (2008)
  10. Dascalu, M.: *Analyzing Discourse and Text Complexity for Learning and Collaborating*, Studies in Computational Intelligence, Vol. 534. Springer, Switzerland (2014)
  11. Trausan-Matu, S., Dascalu, M., Rebedea, T.: A system for the automatic analysis of Computer-Supported Collaborative Learning chats. In: ICALT 2012, pp. 95–99. IEEE, Rome, Italy (2012)
  12. Trausan-Matu, S., Dascalu, M., Dessus, P.: Textual Complexity and Discourse Structure in Computer-Supported Collaborative Learning. In: ITS 2012, Vol. LNCS 7315, pp. 352–357. Springer, Chania, Greece (2012)
  13. Bereiter, C.: *Education and Mind in the Knowledge Age*. Lawrence Erlbaum Associates, Mahwah, NJ (2002)
  14. Scardamalia, M.: Collective cognitive responsibility for the advancement of knowledge. In: Smith, B., Bereiter, C. (eds.) *Liberal Education in a Knowledge Society*, pp. 67–98. Open Court Publishing, Chicago (2002)
  15. Trausan-Matu, S.: Computer Support for Creativity in Small Groups using chats. *Annals of the Academy of Romanian Scientists, Series on Science and Technology of Information*, 3(2), 81–90 (2010)
  16. Trausan-Matu, S., Rebedea, T.: Polyphonic Inter-Animation of Voices in VMT. In: Stahl, G. (ed.) *Studying Virtual Math Teams*, pp. 451–473. Springer, Boston, MA (2009)
  17. Trausan-Matu, S., Stahl, G.: Polyphonic inter-animation of voices in chats. In: CSCL'07 Workshop on Chat Analysis in Virtual Math Teams, pp. 12. ISLS, New Brunswick, NJ (2007)
  18. Trausan-Matu, S.: Automatic Support for the Analysis of Online Collaborative Learning Chat Conversations. In: 3rd Int. Conf. on Hybrid Learning, Vol. LNCS 6248, pp. 383–394. Springer, Beijing, China (2010)